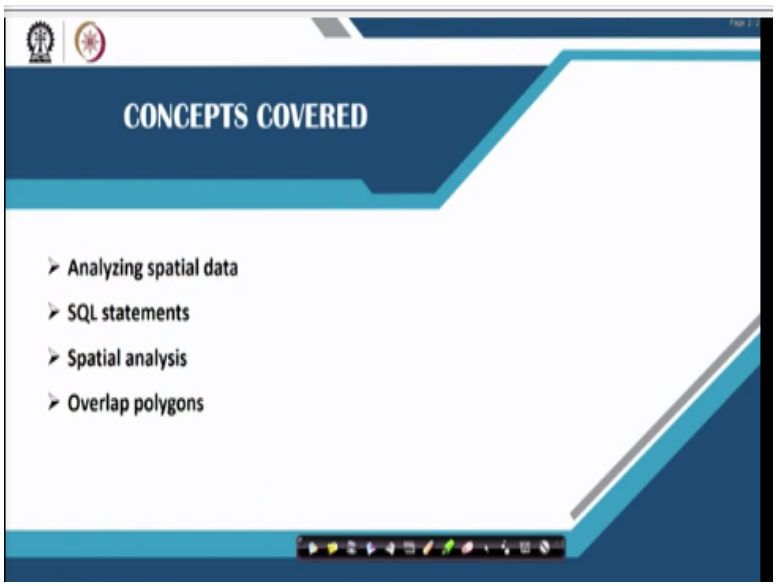
Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology-Kharagpur

Module-09
Lecture-42
Basic Spatial Analysis

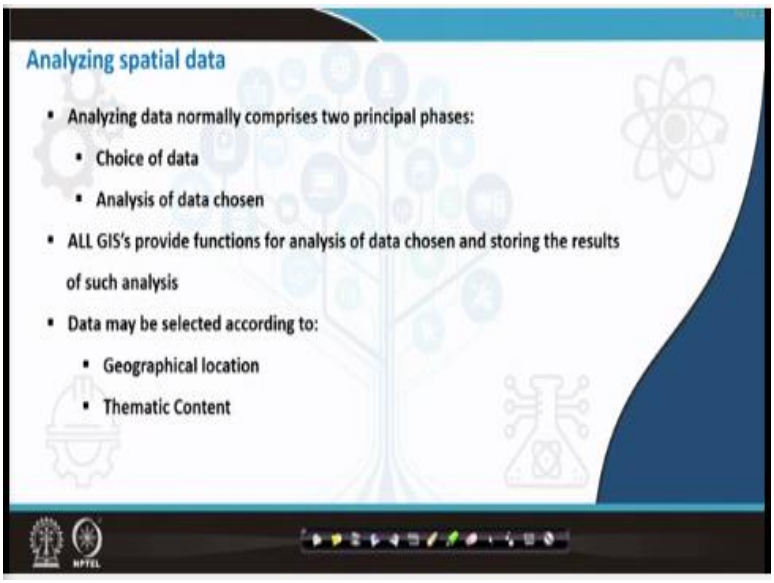
Hello namaste, welcome back to the course on geographic informations system. In this particular session we will let us understand the basic spatial analysis of geographic information system, some part of it I have already spoken. But now this is more into examples and how we actually use these basic spatial tools in our analysis into maybe into urban planning or urban data extraction or any of those aspects ok.

(Refer Slide Time: 00:55)



So now in this class as I said we would look at some analyzing spatial data some part of SQL statements but I am not going into completely into SQL statements. We would look at spatial analysis and some of, for example polygon versus polygon overlaps, polygon on point overlaps etc and different aspects of it.

(Refer Slide Time: 01:16)

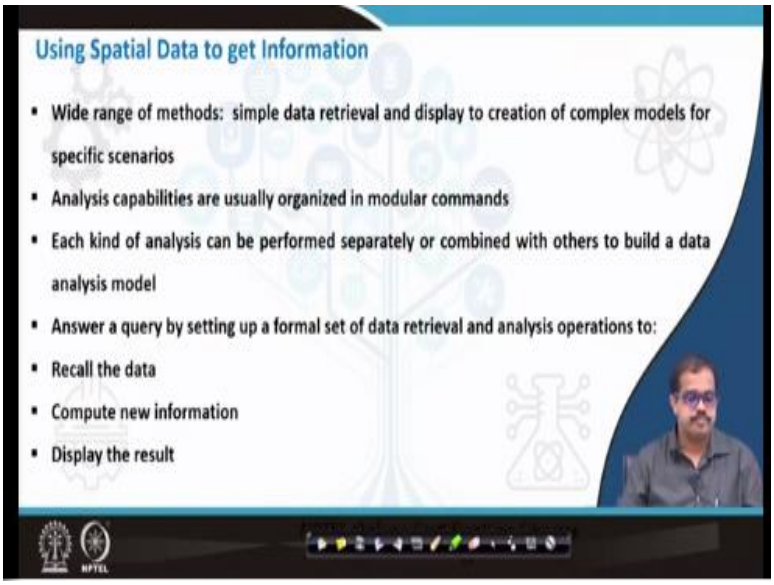


Now, once the first thing is analyzing a spatial data when we have to analyze a spatial data, normally it has done in 2 phases ok. One is a choice of data then analysis of the data chosen, I have well spoken a lot of things about the choice of data, what kind of data has to be acquired, how it has to be acquired. In my last class I even spoke about remote sensing, so you have a different choice of data the way the data has to be connected, the data has to be considered how the data has to be brought into the system is completely dependent on the user.

Then the second thing is analysis of the data chosen, so that is where the GIS is the tools GIS is as a software, as a tool is extremely important in understanding. So all softwares of geographic information system or entire system functions for the analysis of data chosen and storing the results of such analysis. Which means to say that we GIS is more interested in terms of analysis of the data which means the tools that actually developed is in terms of analysis of the data ok.

So when we are looking at data, data may be selected based on 2 things, one is your geographic location other one is the thematic content which we have seen in previous week.

(Refer Slide Time: 02:38)



And when we are looking at how do you get a information from it is data. Now data has to be then converted as a information which means there needs to be certain processing. So you have huge number of methods very simple method is data retrieval and display of it or creation of complex models for specific scenarios ok, you can either use this data to model the same scenario, you can just use the data to represent a scenario.

So it is upon the user to what kind of analysis that particular person is trying to use the data can be used basic analysis and also for the very advanced analysis. So when you look at capabilities, these are usually organized into modular commands ok. Normally it has more of commands but now it has evolved over a period of time. Now, it has more you have a lot of improvements in terms of open source GIS.

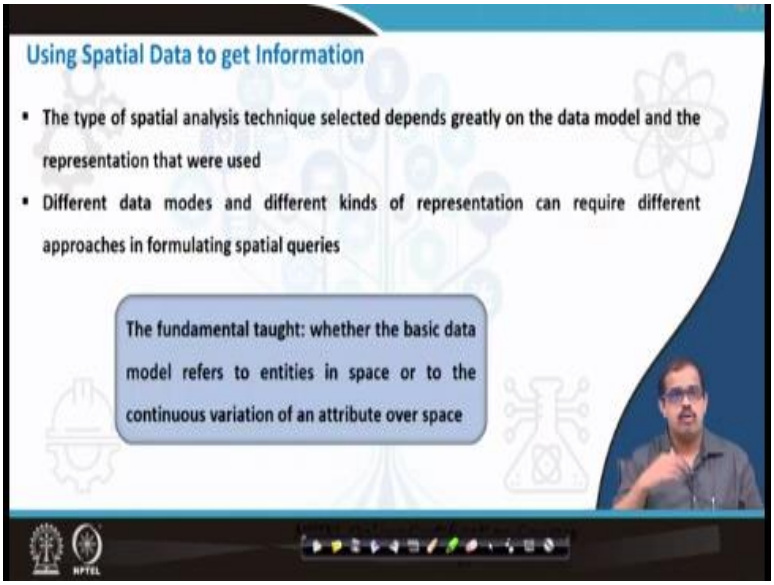
So you can any of your software, GIS softwares has just dropped down menu. So it can be just modular drop down menus. Each kind of analysis can be performed separately or it can be combined with others to build a data analysis model. So which means to say that, if you want to build a data, you can actually look at every data as a separate model and come to the last part of analysis or combine all the data models.

Then develop a more complex data models to give a complete solution, it is depends on again with to the user the way he or she wants to look at it. So when you are looking at a query, the

very formal set of data retrieval is to either recall the data which means retrieve the data and put it into the use right now. Compute that information which you would have recall the data and then compute the information and finally display the results.

That is why you need to have a query, that is why SQL statements are necessary. You need queries that are actually handling the data, capture the data, it process the data and you have a display of results ok.

(Refer Slide Time: 04:46)

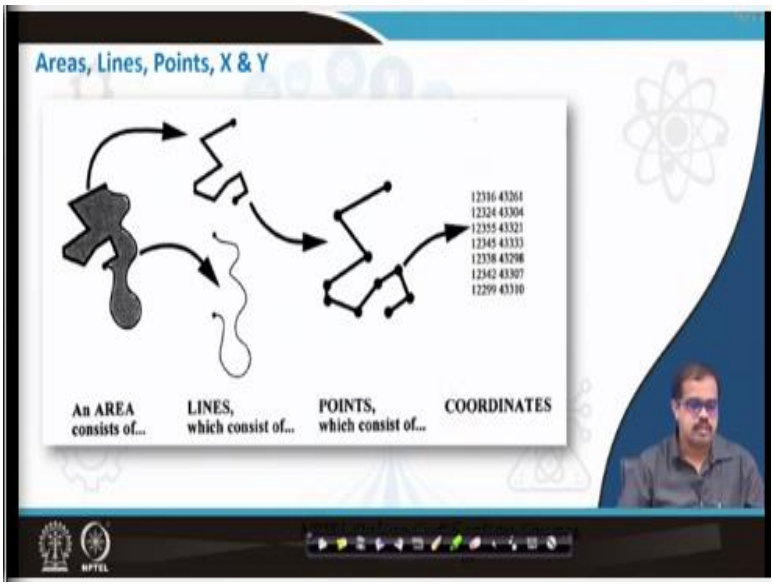


So when you are looking at the spatial data to get information, the first thing that you have to look at is what kind of technique that you have to use which means to say that these techniques are selected depending on greatly on a particular data model. Whether it is a raster data model or a vector data model, what kind of data it is and how the representations are there. So the first thing that you have to see is what kind of spatial analysis technique you would actually taken and what kind of different representations are there in that particular model ok.

So based on that you will be able to process your spatial data to get information. Different data models and different kinds of representation can require different approaches in formulating spatial queries ok. When you are looking at this aspect, whether it is a basic data model or a it is advanced data model, it refers to entities in space or continuous variation of attributes over space.

So that is what you have to understand, it is either it is entities ok or continuous attributes. So it is everything is over space and you are actually making the real world in spatial quantities.

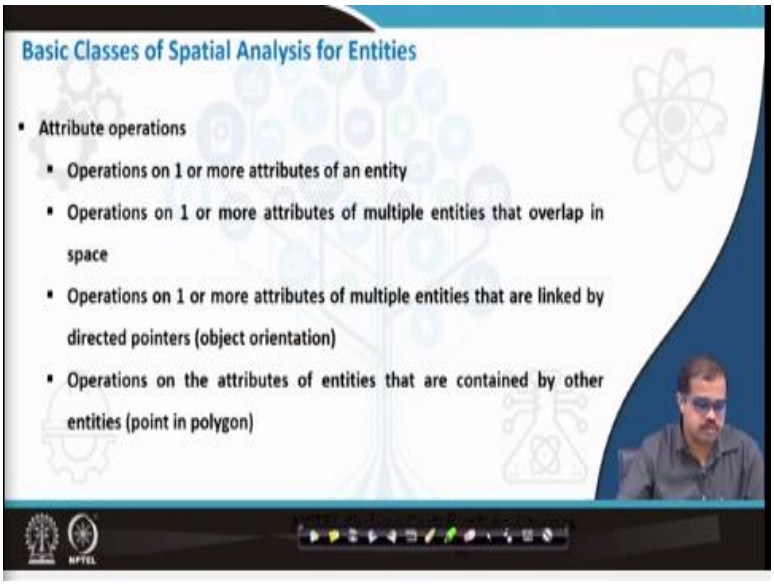
(Refer Slide Time: 06:03)



So when this I have already spoken, so normally when you are looking at the vector data model it is in the form of areas, areas are consist of number of lines. And each line consists of number of points, and each point is representing a coordinate here. So each of these are different coordinates in this particular system, so when you are looking at this particular data model. So if you have coordinates, you can develop the area, if you have areas you can break down into different coordinates.

So the first thing is what kind of data representation model is that, if this is a representation model, what kind of analysis can be done. So that I am just trying to give you an example of how a vector data model actually is and how the information has to be sort out.

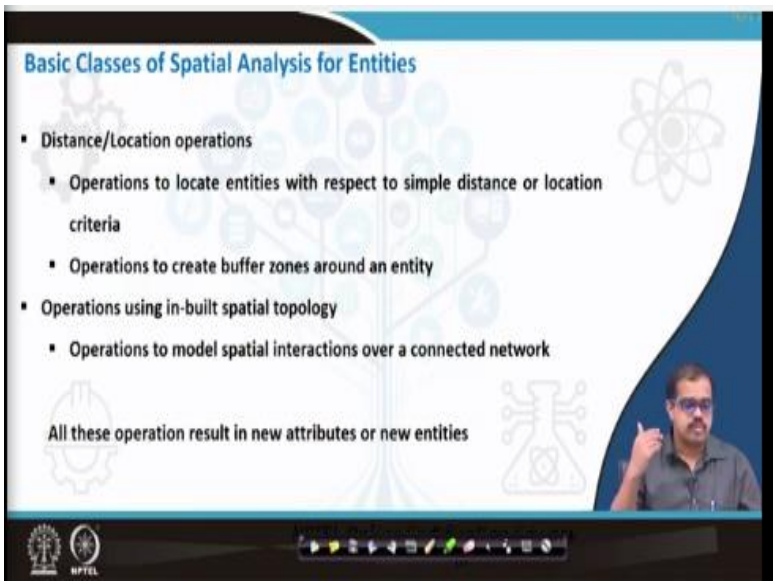
(Refer Slide Time: 06:48)



Then when you look at basic classes of spatial analysis for entities when I say entities, these are separate objects there are not continuous it is discontinuous object. So attribute operations can be operations on one or more attributes of the entity, you can use attributes of the entity for doing the operations or it can be one or more attributes ok. It can be a single attributes or it can be add one or more attributes that actually overlap in space.

It can be operations on one or more attributes of multiple entities that are linked by directed pointers, which is object orientation. It can be operations on attributes of the entities that contained by other entities or point to polygon, it can be one entity is point other entity can be polygon. So looking at this aspect also can be done, so these are the basic classes of spatial analysis that you can do ok.

(Refer Slide Time: 07:50)



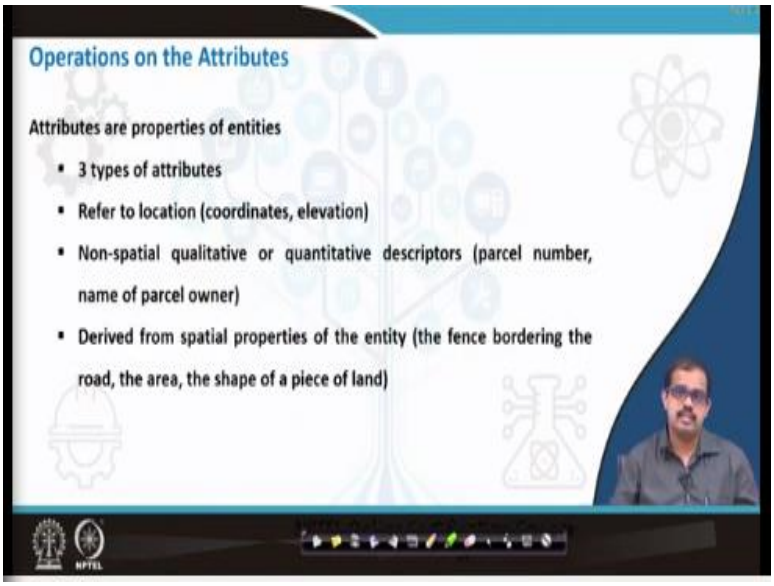
So when you are looking at it some other examples of such basic operations are distance or location operations when you are looking at this operations locate entities with respect to simple distances or for example, if you have built up the entire vector data model of a particular campus. So now you have to find out the distance or location aspect which means to say that you are trying to find out from point A to point B what is a particular distance.

So that is also a basic spatial analysis where you are looking at 2 points or a 2 or a line that is starting and ending ok which is nothing but set of points. Now it can be also to create a buffer zone around an entity. For example, you have a lake in your campus, so you want to create a buffer zone, so that you have that lake intact. So which means to say that buffer analysis is also a speed basic spatial analysis.

Then operations using in to build that topology, I hope you guys understand what do you mean by topology, I have spoken it out in the last week. So if you have to build a basic spatial topology into a connected network then it is also a basic class of analysis. So all these whatever the kind of analysis that you do on the particular data that has been acquired it actually creates a new attributes ok.

If you are adding a buffer it creates another attribute of a buffer and or a new entities, this particular new entity can be an information can be another attribute also ok. So keep this in mind, so it can be both information and an attribute.

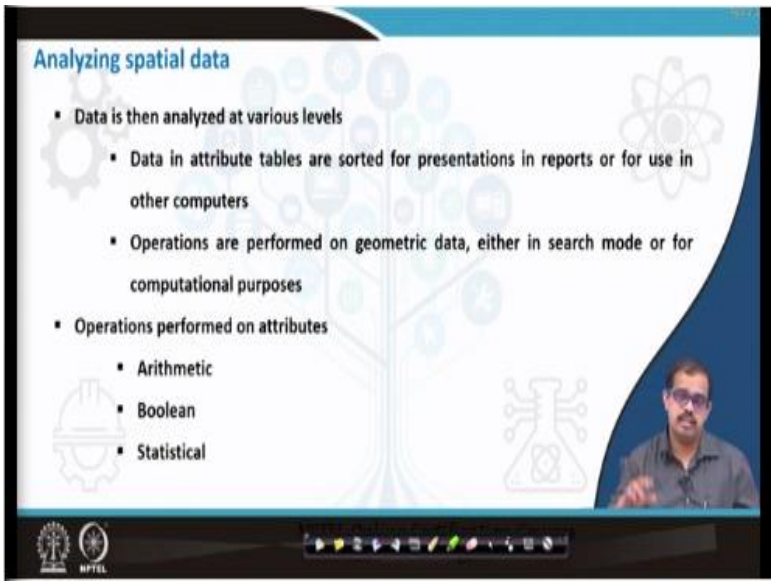
(Refer Slide Time: 09:35)



Now, once you are looking at the other now we have looked at how the data can be easily analyzed. But the next set is you can also look at spatial analysis on attributes ok. When you are looking at attribute, there are 3 types of attributes ok. One refers to a location, you are looking at only location information either a coordinates or elevation of that particular data, so that becomes the first kind of attribute.

The second one is the non spatial qualitative or quantitative descriptors, so that is a second type of attribute. Third one is the attribute that is actually derived by the spatial products of an entity ok. So if you have different entries, you would have done some special operations you would have got another entity, there is a third kind of attribute. So using these 3 different properties there are operations that can be done on each of this data.

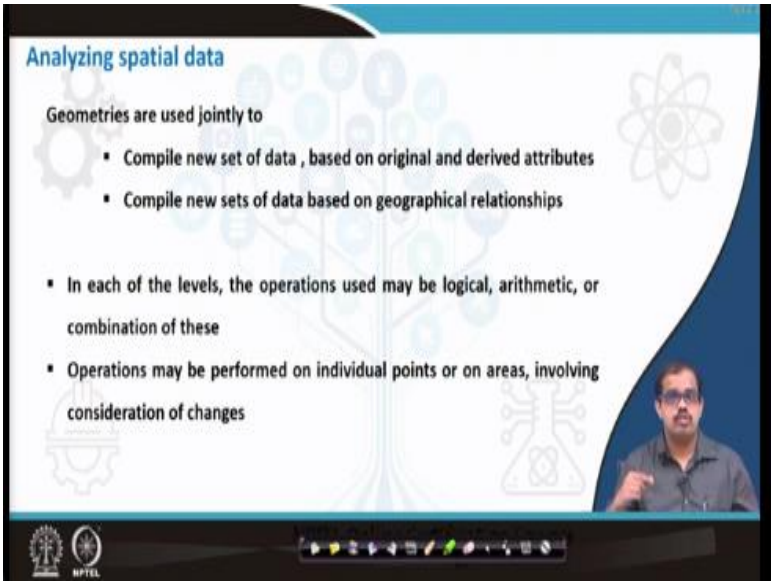
(Refer Slide Time: 10:28)



It can be your arithmetic operation, it can be a Boolean operation, it can be a statistical operation, it can be any kind of operations that you may infer to in getting the outputs. And when you are looking at this spatial data, data is normally analyzed at attribute tables or sorted out for presentation in reports or for the use for in other computers. Operations are performed on geometric data either in search mode or in for the computational purposes.

So when you are looking at operations, I spoke about operations performed. So when you are speaking about operations, it can be either arithmetic operations, Boolean operations or a statistical operations when we say operations.

(Refer Slide Time: 11:14)

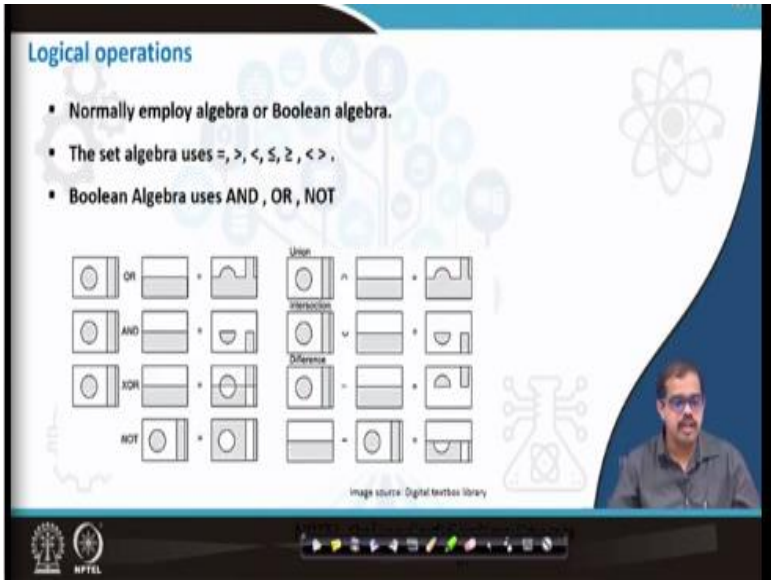


So I am very specific about operations, operations between attributes, operations between attribute and the created information, so that is all, so that is called as an operation. So when I am looking at those operations, it is either arithmetic, it is Boolean or it is either arithmetic Boolean or statistical ok. So when you are looking at this, the very important aspect that I spoke in the previous slide here is a geometric data.

When I say geometric data, it is associated with the geometry, shape of a particular area, object or any of the aspects that we have captured. So when I look at geometries, geometries are used jointly to compile a new set of data based on original or derived attributes, it can be derived attributes. But geometry can be also original, compile new set of data based on geographical relationship, please understand this, I am speaking about both relationships and attributes.

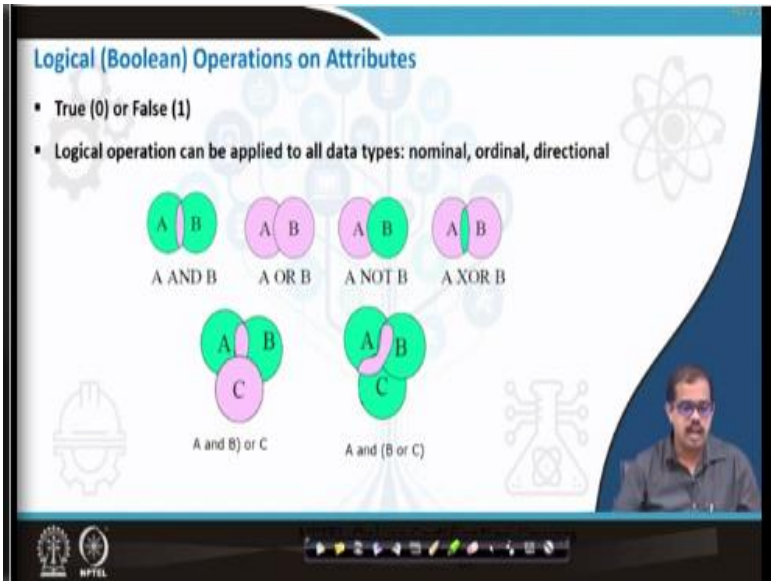
So geometries can be used to have compile or create both attributes compile a new set of data based on geographic location. So geometry forms a very basic part of processing of your data ok. In each of these levels operations use maybe logical arithmetic or combinations, it can be anything, it can be use an logical operation, arithmetic operation or use statistical operation. So or combine all of these to use that is dependent on a user and what kind of analysis is trying to. So operations may be performed on individual points on or an areas involving considering the changes, ok.

(Refer Slide Time: 12:59)



So when you are looking at logical operations normally this employs algebra or a Boolean algebra. So most of you know what do you mean by a Boolean algebra, so let us say that it is equal to, less than, greater than, greater than equal to, less than equal to. So all of these are the things that have been operators that are being used. And Boolean algebra when I say it is AND, OR, NOT, XOR, XNOR all of these come under the Boolean algebra ok.

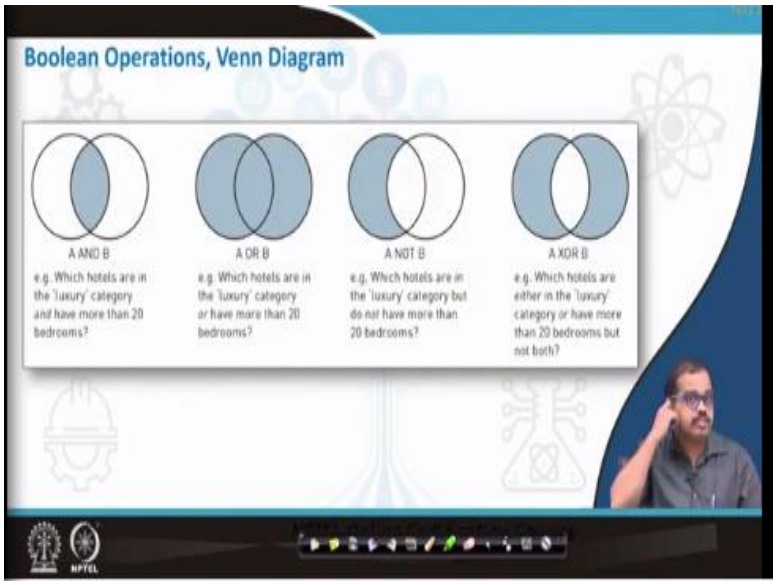
(Refer Slide Time: 13:34)



And to just give you an example here, this A and B which is actually representing both A and B, ok and this is the one that is left out. If it is A or B ok either of the sets A or B, if it is A not to B, it is the one that is B and not A ok, A XOR B exclusively R ok, A and B or C. So these are different operations, logical operations or the Boolean operations that you can do in terms of analysis.

So when you are looking at logical operations, it can be applied for all kinds of data, it can be a nominal data, it can be ordinal data, it can be a directional data, any kind of data that is there, logical operations can be applied.

(Refer Slide Time: 14:23)

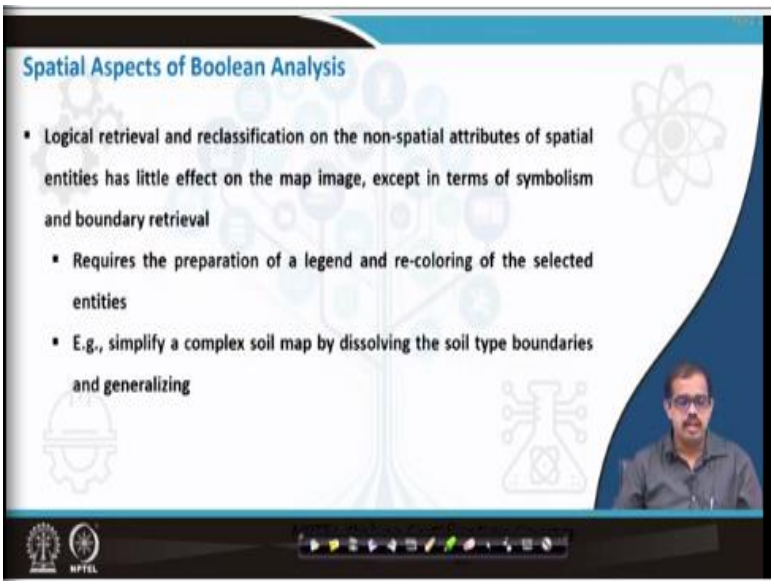


And just to give you an example of all of these if it is A and B ok, for example, if I want to look at a query, which hotels are in the luxury category and have more than 20 bedrooms. Which means I should filter out the data which is in the hotels with luxury category and also having 20 rooms, more than 20 rooms, so that comes here ok.

Then, if it is A or B which are the hotels are in the luxury category, or having which means either a luxury category or having 20 rooms, so not both ok, it can never be both. Then which are the hotels are in the luxury category but do not have more than 20 rooms. If someone wants to have bigger rooms and are trying to look at a smaller properties for their stay, they would actually look at the luxury hotels but in the number of rooms are smaller.

So that it is much quieter ok, which hotels for example, if you are looking at operations like A XOR B, which hotels are either in luxury category or have more than 20 rooms but not both, then you have to look at XOR B. So these are different Boolean operations that you can look at ok.

(Refer Slide Time: 15:42)

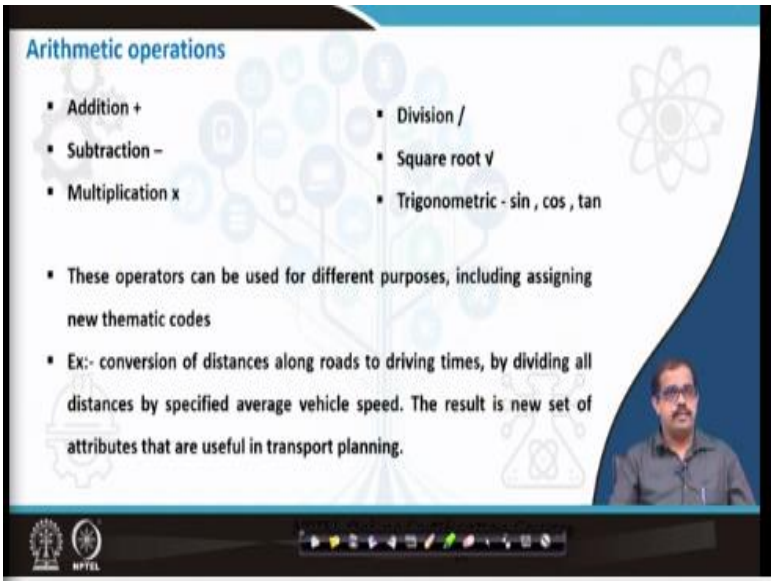


So when you are looking at the spatial analysis of any of the Boolean operations, so logical retrieval and reclassification can be easily done on a non special attributes of spatial entities ok. Non spatial attributes of the spatial entities, but these do not affect the map image much ok except in terms of symbolism, and boundary retrieval, ok. Otherwise it does not affect much ok, just putting your color, putting a shape all this or whatever you represent it in a form of a different colors etc.

So all those things are only affected but it does not really affect the map, ok. But it requires a preparation of a legend and recoloring of a selected entities, simplifies a complex map by dissolving soil type boundaries. For example if there is a map which is representing a soil type boundaries. So if you are trying to do this logical retrieval and reclassification, so that you do not need, so many types of soils you are looking at a more broader version.

So to just do an retrieval of whatever the query you have, so you probably the Boolean analysis can be used, that is what I am trying to say in this particular slide.

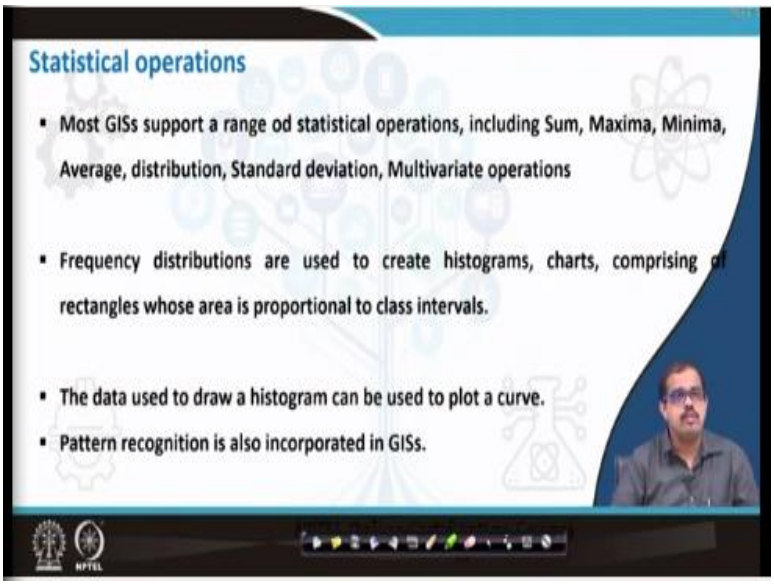
(Refer Slide Time: 16:57)



And when you are looking at the next set is the arithmetic operations, that is the addition, subtraction, multiplication, division, square root, trigonometric etc. So these operations can be used for different purposes is including assigning new thematic codes. That is for example conversion of distance along the roads to driving times by dividing all distances by specified average vehicle. The result is a new set of attributes that are useful in transport planning.

So this is one example or for example, if you are for example, there are 2 urban area maps ok, I have extracted only urban area. So now I have to see what is the difference, so it is as simple as this. So you have older map just make it as the urban area as 0 and the newer map the urban area is 1. So multiply both where there is 0 the entire region is actually awarded of ok. And the new urban area which is actually 1 is actually multiplied here ok, you can extract that particular data. So that is example of arithmetic operations that you can do on a map or an image ok.

(Refer Slide Time: 18:05)

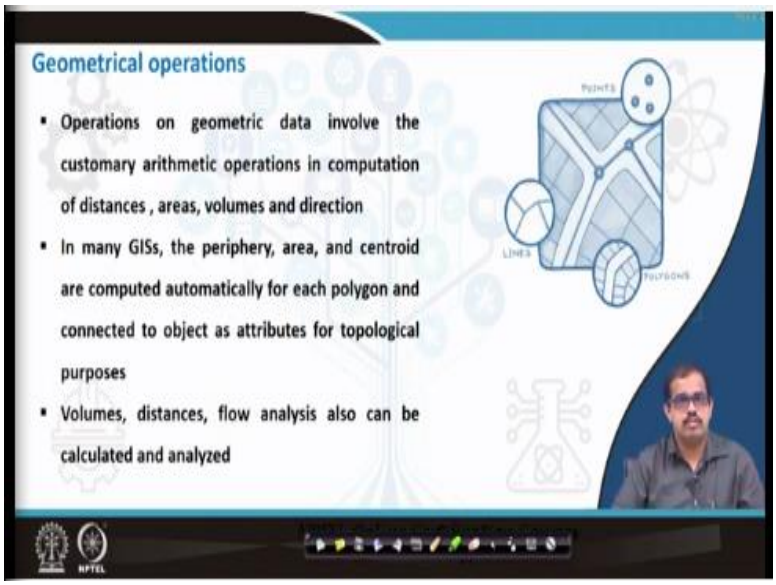


Then you have a statistical operation these are the ones which are considered for advanced operations and GIS packages ok. When I look at statistical operations all kinds of statistics, so I have just mentioned it is some maximum, mean, average distribution deviation, multivariate operations. So other than this whether it is interpolation, extrapolation and any kind of statistical operations can be done using a GIS software.

Most of the GIS software supports most of the statistical operator and when you are looking at this frequency distributions are one of the very important aspects when you are looking at spatial data. So you can create histograms charts comprising of rectangles, whose area is actually proportional to the class intervals also, if you are looking at class data. And the data use to draw or arrange data or the data used to draw in histogram can be even a plotted in to a curve.

I am trying to say that whether you are using a statistical operations or whether you are putting it in a form of an image. Everything can be done using an GIS software ok, any kind of analysis. And one of the very important aspects of GIS I would give you an case study of or of many cities of India at the end of probably end of this course. Where I have I would show you how the pattern recognition is applied on the cities in that is a just a part of how GIS can be effectively used for analyzing the situation ok.

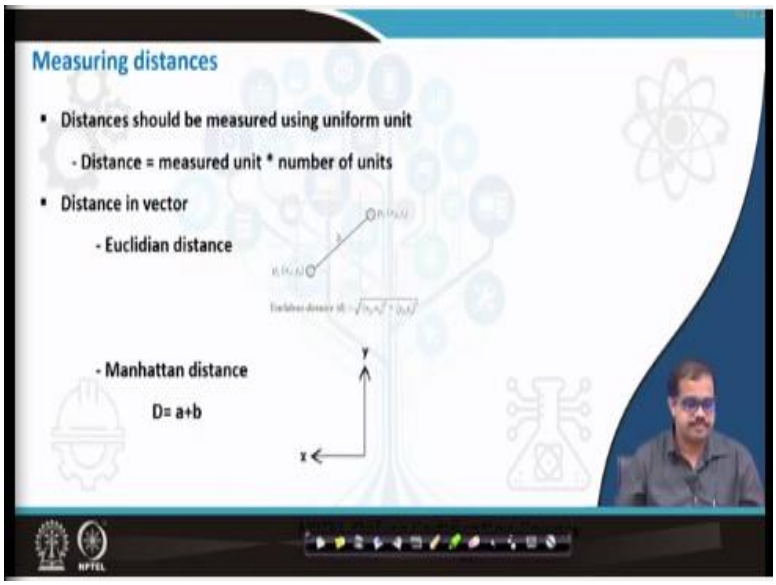
(Refer Slide Time: 19:44)



So if geometrical operations are the ones the order of the day, so any geometric data involves a customary arithmetic operations and computation of distance, area, volumes and directions. So when you are looking at GIS the whether you are looking at a centroid or a periphery area or you are trying to look at how each polygons are computed automatically, how an object is connected with an attribute or for topological operations.

So all of these are done through geometrical operations or what kind of analysis whether it is a line, polygon or a point. So all of these calculated using the geometrical operations. And when you are looking at volume, distance flow analysis all of these are calculated and can be easily analyzed using geometrical operations. So that is why you have a class of operations which is called geometrical operations.

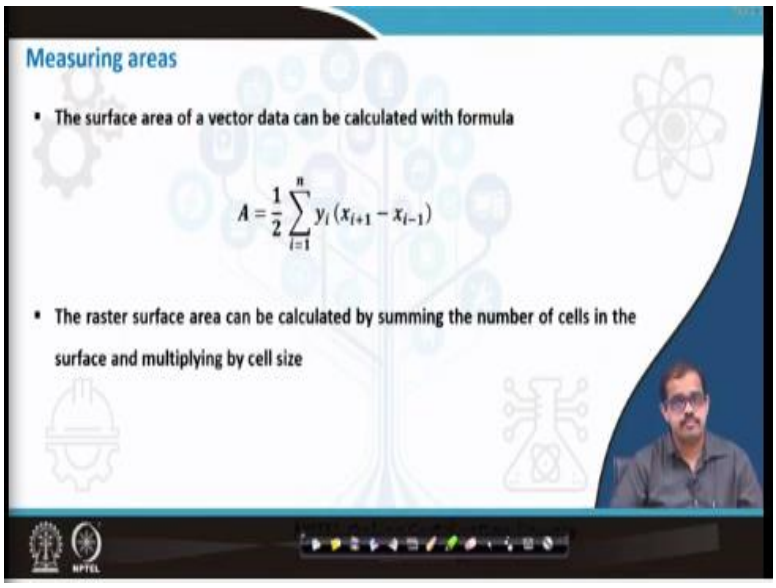
(Refer Slide Time: 20:34)



For example, let us take some of them, so first one is measuring distance. And if you want to measure a distance, for example, you are measuring the distance from this point to here, it is as good as calculating a Euclidian distance. When I say Euclidian distance, if it is x 1 y 1 and x 2 y 2, so it is x 2 - x 1 the whole square + y 2 - y 1 the whole square under the square root.

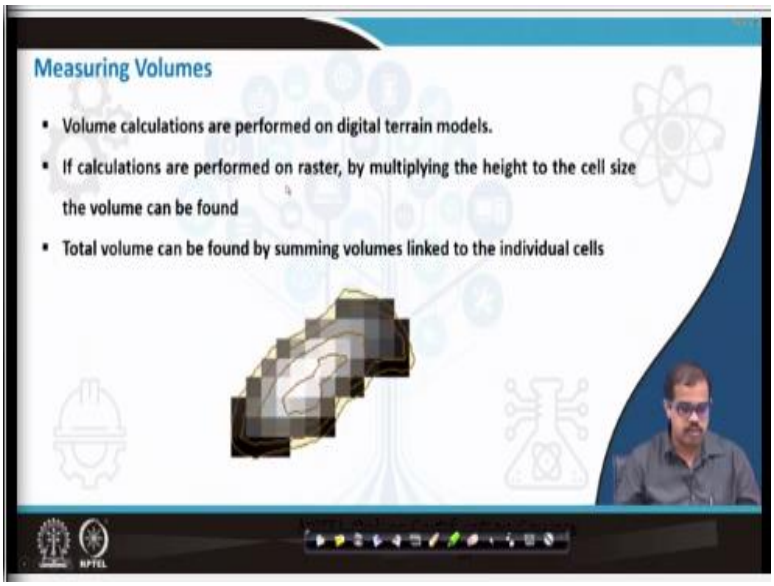
So this is a distance of a vector ok which has to be calculated. So if it has to be calculated on a map at it is distance is equal to measured unit into number of units as simple as this. And if you are calculating a Manhattan distance if it is given by a + b, which means distance on a and b. So that is how you calculate a distance of a vector.

(Refer Slide Time: 21:23)



And if you are measuring the surface area, so it can be just calculated using this particular formula which means a raster surface can be calculated by summing the number of cells in a particular surface multiplied by the cell size ok. If it is a spatial raster data, then as simple as this probably when I say cell size, it is the area that each cell is actually covering. So that is what this particular formula explains, so that is how you are areas are calculated.

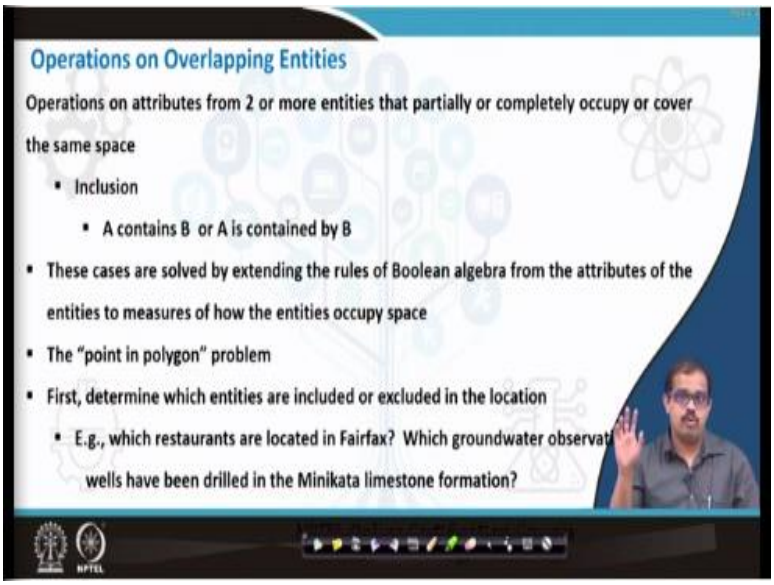
(Refer Slide Time: 21:55)



And that particular unit, so most of them once they calculate the area they are worried about what is a units. So unit is actually represented by the terms that particular map is prepared, if the unit has kilometers, it has an meter or kilometer square, if it is in meters it is meter square. So depending on what kind of unit the unit is assigned to your final calculations. Then measuring volumes, so total volumes can be by summing the volumes linked to the each individual set.

So now find out the volume of each individual cell here then sum all of those to get a sum of the volumes. So if you are trying to do it by a raster, so you have to multiply the height of the cell size and the volume that can be found in each of these cell size as simple as that ok.

(Refer Slide Time: 22:46)

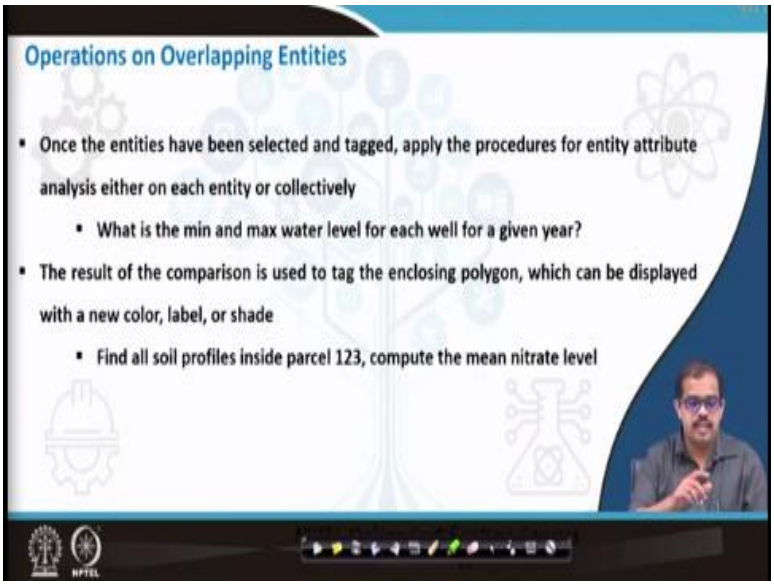


So what am I speaking theoretically is looks simple but when you are actually putting it on in a software it is more complicated ok. So when you are looking at the next set of options or analysis that you can do is operation of overlapping entities. The first kind of operations that you can do is point and polygon or first operation that we look at this point in polygon.

So when you are looking at operations on attributes from 2 or more entities, that partially or completely occupy or cover the same space is inclusion. A contains B or A is contained by B, these cases are solved by extending the rules of Boolean algebra from attributes of the entities to measure of how the entities occupy a particular space. The point in polygon problem can be easily understood here.

So first thing is which restaurants are located in Fairfax, which groundwater observation wells have been drilled in limestone formation. So these are the thing queries that can be easily found when you are having a point and polygon problem.

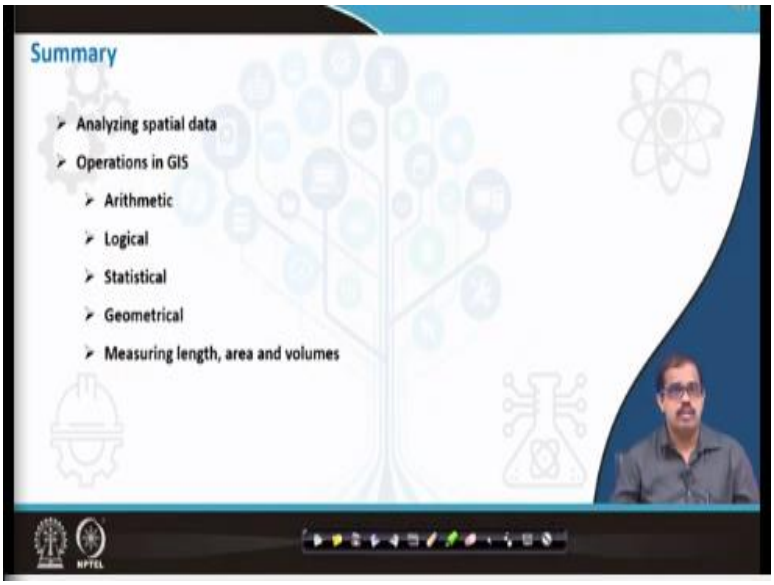
(Refer Slide Time: 23:56)



I will go back to this in the next slide. Once entries have been selected and tag we have to look at entity analysis either on each entity or collectively. You can take each entity and look at the analysis part or collectively of this, this I have spoke before also. So for example, what is a minimum and the maximum water level for each well. So when you are looking at each well, it is each entity, if you are looking at a collectively the entire area wells in the area in a given year then it is the collectively or looking at the entire area.

The result of comparison is used to tag or enclosed polygon which can be displayed in a new color. So now you have a set in one particular layout another set in another particular layout another set in another layout. Now, you will use different colors for each of these set to represent your output ok. And it may be looking at different soils, different amount of water that is present in the well whatever that this is that you are trying to do ok.

(Refer Slide Time: 24:56)



So in summary, we have understood how the basic spatial analysis is done, we have looked at operations in GIS. For example what do you mean by an arithmetic operations, what is a logical operation, a statistical operation, a geometrical operation and measuring length areas and volume, so all of these we have looked at as of now. So in the next class we will continue this, we will look at more of different basic spatial analysis.

And once we have understood that the finally we would look at that one spatial analysis and would stop with what are the limitations of what GIS and GIS software are at today's context and how it can be improved. So that is what we are trying to do in this entire week, so till the next class have a nice time, thank you.